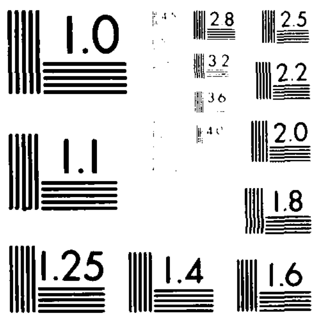
AD-A088 677   OKLAHOMA UNIV NORMAN DECISION PROCESSES LAB   F/G 5/10
HYPOTHESIS GENERATION IN AN AUTOMOBILE MALFUNCTION INFERENCE TA-- TC(U)
FEB 80   T MEHLE   N00014-77-C-0615
UNCLASSIFIED   TR-25-2-80   NL

1 OF 1
AD
A088677

END
DATE
FILMED
10-80
DTIC



MICROCOPY RESOLUTION TEST CHART

DECISION PROCESSES LABORATORY
UNIVERSITY OF OKLAHOMA

OFFICE OF NAVAL RESEARCH
CONTRACT NUMBER N00014-77-C-0615
WORK UNIT NUMBER NR197-040

Hypothesis generation in an automobile malfunction inference task

Thomas Mehle

TR 25-2-80 February 1980

prepared for

Office of Naval Research
Engineering Psychology Programs
Contract Number N00014-77-C-0615
Work Unit Number NR 197-040

Decision Processes Laboratory
Department of Psychology
University of Oklahoma
Norman, Oklahoma 73019

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: TR-25-2-80	2. GOVT ACCESSION NO.: AD-A088 677
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): HYPOTHESIS GENERATION IN AN AUTOMOBILE MALFUNCTION INFERENCE TASK	5. TYPE OF REPORT & PERIOD COVERED: Technical Aug. 1979-Aug. 1980
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Thomas Mehle	8. CONTRACT OR GRANT NUMBER(s): N00014-77-C-0615
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Decision Processes Laboratory, Department of Psychology, University of Oklahoma, Norman, Oklahoma 73019	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 197-040
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Engineering Psychology Programs-Code 455, 800 N. Quincy Street, Arlington, Virginia 22217	12. REPORT DATE: 25 February 1980
	13. NUMBER OF PAGES: 42
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

1) decision theory
2) hypothesis generation
3) verbal protocol analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Expert and novice subjects generated hypotheses in an automobile trouble-shooting inference task. Data collected included subjects' verbal protocols during the inference tasks and subjects' estimates of the probabilities of their generated sets of hypotheses. Analyses indicated that both expert and

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

410 118

(continued from Block 20)

novice subjects had difficulty generating complete sets of hypotheses and were overconfident in their subjective estimates of the probabilities of generated hypotheses.

Abstract

Expert and novice subjects generated hypotheses in an automobile trouble-shooting inference task. Data collected included subjects' verbal protocols during the inference tasks and subjects' estimates of the probabilities of their generated sets of hypotheses. Analyses indicated that both expert and novice subjects had difficulty generating complete sets of hypotheses and were overconfident in their subjective estimates of the probabilities of the generated hypotheses.

Hypothesis Generation in an Automobile Malfunction Inference Task

Hypothesis generation can be a critical component of decision making in problems for which hypotheses concerning possible states of the world are not obvious. Such problems constitute an important class; they are common in the realms of scientific investigations, mechanical and electronic trouble-shooting, medicine and societal decision making. In these problems, poor hypothesis generation may lead to neglecting possible states of the world in subsequent analysis, thus degrading the entire decision-making process. The purpose of the research described here was to examine hypothesis generation and assessment in the context of automotive trouble-shooting.

Hypothesis generation and hypothesis assessment are not necessarily independent processes; they can interact through-out the problem-solving process. A retrieved hypothesis must be considered somewhat plausible initially if it is to be entertained. If for some reason all hypotheses are rendered implausible, a decision maker is likely to resume retrieval activities.

Recent events in the nuclear power industry serve to act as an example of how hypothesis generation, hypothesis assessment and decision making can interact. In making decisions concerning the operation of nuclear power plants, it is important that decision makers generate all hypotheses concerning safety device failures; the alternative is an overestimate of the probability that the nuclear plant will operate safely. As an illustration, prior to the incident at the nuclear power plant on Three Mile Island, operators had closed all three auxiliary feedwater pumps. This action was in violation of Nuclear Regulatory Commission rules and made the emergency cooling system inoperative. It is likely that the decision to permit the operation of the Three Mile Island power plant was based in part on an analysis that did not anticipate this state of the world.

Despite the crucial importance of hypothesis generation in many contexts, it has received little attention until recently. An early exception was Hanson (1961), who noted that the importance of the hypothesis generation process was alluded to by Aristotle (Prior Analytics II, 25). Hanson approached hypothesis generation on philosphical grounds, arguing that the process by which a hypothesis is generated as a plausible alternative worth entertaining is logically distinct from the process by which hypotheses are evaluated. Hanson examined the process by investigating the historical accounts of hypothesis generation by exceptional scientists, notably Kepler. Hanson's description of hypothesis generation was as a three-step process. The first step was the decision maker becoming aware of an anomoly in the data; the anomoly was the stimulus for hypothesis generation. Secondly, a hypothesis was generated and lastly, it was incorporated into an organized system of concepts. In other words, it is detection of an anomoly in the data which acts as the stimulus for further hypothesis generation.

Churchman and Buchanan (1969) characterized hypothesis generation, which they termed an "inductive process," as a two-component system. In their model, "H" is a hypothesis, "D" is the data to be explained and "E" is the problem context. The two components are: 1) Find an H which satisfies the schema: D because H and E. 2) Determine if H satisfies a "satisfactoriness" criterion.

In investigating hypothesis generation in mass spectrometry problems, Churchman and Buchanan expanded these two into eight steps, which were incorporated in a computer program. Briefly, the steps were: 1) collecting the data, 2) interpreting the data, 3) selecting the general class of plausible hypotheses, 4) limiting the number of hypotheses through testing, 5) generating specific plausible hypotheses, 6) making predictions, 7) evaluating the satisfactoriness of the hypotheses that have been generated and 8) recycling if no satisfactory hypotheses were generated.

Churchman and Buchanan's term "satisfactoriness" can be identified with evaluation of hypothesis plausibility; their seventh step is analogous to Hanson's first step. Churchman and Buchanan's orientation in examining hypothesis generation was primarily philosophical; one of the major conclusions of their paper was that inductive systems (i.e. hypothesis generation processes) in the empirical sciences are not even approximately rational.

Although their primary concern was scientific inference, Gerwin's (1974) and Gerwin and Newsted's (1977) discussion of hypothesis generation is relevant in a broader context. Gerwin (1974) pointed out that Hanson's (1961) view of hypothesis generation is closely related to the views of Simon (see Simon's 1978 article for a review of his work). One of Simon's interests has been to explain, model and predict the verbal behavior of subjects instructed to talk aloud while solving problems. Simon has been a proponent of the view that psychological research should examine the specific behavior of individuals rather than aggregates. In a 1975 article, Simon asserted that "diversity of behavior may be hidden under a blanket label...we must avoid blending together in a statistical stew quite diverse problem solving behaviors whose real significance is lost in the averaging process," (p. 288).

The emphasis of Simon and his associate's work has not been hypothesis generation per se, but this process has been touched on in their investigations of the global problem-solving process. Other researchers employing protocol analysis techniques in investigations of problem-solving behavior have frequently addressed hypothesis generation, at least tangentially. The technique of examining verbal protocols has been used to investigate a wide variety of problem-solving activities, for example: computer programming (Brooks, 1977), medical diagnosis (Wortman, 1966, 1970, 1971; Wortman and Kleinmutz, 1973), apartment renting (Payne, 1976; Payne, Braunstein and Carrol, 1978) and chemical engineering thermodynamics (Bhaskar and Simon, 1977).

The use of verbal protocol data in psychological research has recently come under attack. Doubts of critics were summarized by Nisbett and Wilson (1977). Nisbett and Wilson pointed out that in many circumstances, subjects may be unaware of significant cognitive events and simultaneously very confident that their verbalizations are quite complete. Also, subjects may report what they conjecture has gone through their mind rather than actual mental events.

Ericsson and Simon (1978) presented an exhaustive rejoinder to the criticisms of Nisbett and Wilson, and others. They examined the specific conditions under which verbal protocols would and would not represent useful data. Their conclusion was that verbal protocol data are most reliable and interpretable when subjects are given generalized instructions to verbalize and when the experimenter's additional probing is minimal. Because of the reconstructive nature of memory, it is important that subjects verbalize while performing the task, rather than at some later time. Although this debate has probably not been resolved to everyone's satisfaction, the position adopted here is that verbal protocols do represent useful data when the conditions specified by Ericsson and Simon are satisfied. That is, protocol analysis methodology represents a potentially valuable approach to examining human behavior, as a supplement or a precursor to traditional methodology.

In his discussions of real-world problem-solving behaviors, Simon (1979) noted the importance of examining "semantically rich" domains; i.e., problem domains which require area-specific knowledge in addition to general problem-solving skills. An example is trouble-shooting; see Rouse for a review (1978a) and a model (1978b) of the trouble-shooting task. Rouse's (1978b) model showed promise in predicting subjects' problem-solving behavior. The model was based on fuzzy set theory, a collection of concepts which may have further application in modeling the hypothesis-generation process (see Zadeh, 1965, for an introduction to fuzzy set theory). Rouse also investigated the performance of subjects and the utility of a computer aid. Further discussion of computer aids in trouble-shooting tasks can be found in Sacerdoti (1975) and in Hart's (1975) description of a computerized consultant to aid mechanics.

Trouble-shooting tasks were examined in an insightful series of studies by Fischhoff, Slovic and Lichtenstein (1978). They reported that both expert and novice subjects in an automotive trouble-shooting task were quite insensitive to the removal of relevant pathways to possible causes of malfunctions and were overconfident in judging the exhaustiveness of "pruned branches" of fault trees. Their investigations supported an availability hypothesis (Tversky and Kahneman, 1973) as the significant contributor to this overconfidence bias. In a somewhat related context, overconfidence has been reviewed and studied by Slovic and Fischhoff (1977), Fischhoff, Slovic and Lichtenstein (1977) and Lichtenstein, Fischhoff and Phillips (1977).

Fischhoff et al. (1977) reported that the overconfidence bias was quite robust to changes in response mode and that subjects were very willing to back up their biased opinions with cash. They suggested two possible explanations for the observed overconfidence: 1) insufficient acknowledgment of uncertainty in the early stages of inference and 2) insufficient awareness of the reconstructive nature of memory. A robust overconfidence bias was observed in a study of hypothesis-generation by Mehle, Gettys, Manning, Baca and Fisher (1979), who also concluded that the bias may be due in part to the operation of an availability heuristic.

In the first of a series of studies investigating the psychological processes underlying hypothesis generation, Gettys and Fisher (1979) advanced a model of hypothesis generation, proposing that an executive process initiates, directs and terminates highly specific, recursive memory searches for possible hypotheses. They postulated that the stimulus for initiation of hypothesis generation would be low plausibility of the current hypothesis set. From the psychological viewpoint, it would seem that the processes most important to hypothesis generation as a distinct component of problem solving are: 1) retrieval of potential hypotheses from memory, 2) evaluation of candidate hypotheses to determine whether they should be entertained and 3) evaluation of the collection of hypotheses under consideration to determine if retrieval should be terminated or resumed. Of related interest was Fisher, Gettys, Manning, Mehle and Baca's (1979) discussion of memory retrieval involving more than a single datum. Memory retrieval employing multiple retrieval cues has also been studied in a different setting by Shanteau and McClelland (1975).

A primary motivation for the current study was to investigate hypothesis generation in a semantically rich problem-solving domain. The task chosen was automotive trouble-shooting, motivated in part by a desire to examine the behavior of both novice and expert decision makers. The decision was also made to obtain verbal protocol data in addition to the more traditional dependent measure of subjective probability estimates. Verbal protocols would be analyzed in an effort to identify the cognitive mechanisms responsible for behavior observed in previous studies of hypothesis generation behavior. Specifically, in tasks where subjects were asked to infer the major of an unknown undergraduate student from a sample of classes taken by the student, Gettys, Mehle, Baca, Fisher and Manning (1979) reported that subjects generated very impoverished sets of hypotheses.

The study also involved specific instructions for subjects to evaluate the plausibility of their sets of generated hypotheses. This instruction is tantamount to obtaining a subjective estimate of the exhaustiveness of the set of generated hypotheses. As previously noted, the typical result in such assessments is for a large and robust overconfidence bias. It was felt that verbal protocol data would be potentially very useful in identifying the mechanisms responsible for this bias and in determining whether there are differences in this bias between expert and novice subjects. Experts' greater store of semantic knowledge might lessen the bias. Alternately, novices might be aware of their lesser store of knowledge and be relatively less biased.

The present study differs from the Fishhoff et al. (1978) studies of automotive trouble-shooting on a significant dimension. In the Fischoff et al. studies, subjects were provided with hypotheses; subjects in the current study generated their own hypotheses. One possible effect of having subjects generate their own hypotheses might be to increase the overconfidence bias, since subjects' hypothesis sets would be more likely to contain personal favorites.

Method

Subjects

Seven of the twelve subjects participating in this study were male introductory psychology students enrolled as undergraduates at the University of Oklahoma. One of these students had worked as a mechanic in a commercial garage and therefore was classified as an "expert". The remaining six students were classified as "novices". Another five expert subjects were employees of the University Motor Pool; these five subjects were paid a $10 honorarium for their participation in the study. Thus six subjects were classified as novices and six were experts.

Apparatus

Instructions and problems were presented to subjects on a Compucolor 8001, a microcomputer having color graphics capability, manufactured by the Intelligent Systems Corporation, Norcross, GA. Subjects' verbal protocols were recorded on a portable cassette tape recorder for later transcription. Subjects' probability estimates were made with a light-pen attachment on the computer's CRT.

Procedure

Subjects received an extensive introduction to the experimental session. Written instructions presented on the CRT were augmented by the experimenter, who was present during the entire session. The following instructions were among those presented on the CRT: "In this study, you will be concerned with things you normally consider when you first approach a problem. The general situation is:

"Imagine that you receive a telephone call from your spouse when you are at work. The general scene is that your spouse mentions having some car trouble. The computer will elaborate the general scene with descriptions of several specific scenes. Please consider each specific scene to be a new and independent situation."

'Your job will be to describe a list of possible explanations of the car trouble which would explain the situation."

Subjects were instructed to "think aloud" during the experimental session. Verbal protocols were tape recorded with the subjects' knowledge. The descriptions of the five specific problems were inspired in part by reference to an automotive trouble-shooting guide in Milton (1971). The text of the specific stimuli presented to subjects on the five trials is contained in Table 1.

Insert Table 1 about here

For each problem, subjects typed in possible hypotheses on the computer's keyboard while thinking out loud. Subjects were instructed to enter all plausible hypotheses that they would be likely to entertain. When subjects had generated all of their hypotheses for a problem, they estimated the probability that the true cause of the car's problem was among those they had generated. This estimate was obtained by having subjects use a light pen to adjust the colored portion of a line on the computer's CRT. The line had calibration reference marks at 0, 25, 50, 75 and 100 percent of its length. Subjects were instructed to estimate the probability that the true or actual cause of the car's problem was included in their list of generated hypotheses.

Results and Discussion

Protocol Data

Subjects' vocalizations were transcribed verbatim from the tape recordings, separated into thought units (protocols) and consecutively numbered for each subject, preceded by a subject letter code. Thus "A1" would be the reference code for the first protocol produced by the first subject and "B5" would be the code for the fifth protocol produced by the second subject. A protocol is operationally defined as a "meaningful thought unit," as judged by the experimenter (see Ericsson and Simon, 1978).

On initial examination, the most striking feature of the protocol data was the sparseness of verbalizations by subjects, notably experts. Although verbalizations were broken down into protocols primarily to facilitate analyses, a count of the ptorocols does provide a rough measure of verbal fluency. For the entire set of five problems, the median number of protocols generated by expert subjects was only 54 per session; the median for novices was 80.5. The mean number of protocols per problem was 33.4 for novice subjects (range: 7 to 194) and 15.6 for experts (range: 2 to 66). Summary data for the number of protocols is listed in Table 2.

Insert Table 2 about here

The sparseness of experts' verbalizations, in comparison to novices, was unexpected. Perhaps the reason the experts did not verbalize more is that they did not understand the task. This possibility is unlikely in light of subjects' verbal reports during debriefing. Virtually all experts apologized for not saying more, stating that they just could not think of anything more to say.

If subjects understood the task requirements, then two conclusions are possible. Either the verbal protocols failed to track subjects' cognitive processes or the protocols accurately reflect the sparseness of the underlying processes in this task. One factor that might contribute to difficulty in verbalizing is expertise. Simon (1978) reported that vocalizations tend to decrease as subjects become more proficient and responses more automatic.

Another possible factor is career-related skills. Protocol studies in the past have tended to employ verbally fluent subjects, such as physicians (Wortman, 1972) or students enrolled in a chemical engineering course (Bhaskar and Simon, 1977). Such subjects' professional success would be partially a function of verbal fluency; success in auto mechanics is less dependent on verbal skills. This possibility is supported by the observation that Subject F generated 244 protocols, more than five times the average of 45 generated by the other five experts. Subject F was the only expert subject that was also a college student. Similarly, but without any apparent reason, one subject stood out from the novice group. Subject D generated 539 protocols versus an average of 92.6 for the other five novices.

A possible contributor to the low frequency of subjects' vocalizations during hypothesis generation was the intrinsic nature of the task. Hypothesis generation is basically a one-step task. Other investigators have generally studied multi-step tasks, such as the Tower of Hanoi problem, the missionaries and cannibals problem and their isomorphs (Simon, 1975; 1979). In such explicitly multi-step tasks, subjects have numerous opportunities to verbalize as they work through all of the component actions. Perhaps the protocols accurately reflect a relatively simple and unelaborated hypothesis generation process for the typical subject. On the other hand, there may be complex mental events associated with hypothesis generation which simply can not be "tracked" by verbal protocols.

An examination of the verbal protocols did not reveal any major differences in content among the subjects. Therefore, it was decided to concentrate on the protocol data for subjects D and F. These two subjects were the most verbal members of their respective groups and their protocol data contained all strategies and processes identified in the protocol data as a whole. This approach should not seriously compromise the analysis, since the general motivation was to identify strategies and processes, rather than to establish any as frequent or universal.

Novice subject D characteristically generated hypotheses that were subsequently ruled out as inconsistent with the data. For example, for Problem 4 (see Table 1):

D4: Had a recent tune up,

D5: So there's no problem with the points.

D6: Two years old,

D7: So, there can't be a lot of problem with all the gears.

D14: New car,

D15: So, that leaves out the mechanical.

D34: Starts fine,

D35: So, there's no problem with the electrical works at all.

Also, for Problem 1:

D75: Wheel balance?

D76: No, that's nothing to do with car starting.

The preceding protocols provide direct evidence for the existence of a "consistency checking" process during hypothesis generation, also investigated by Fisher, Gettys, Manning, Mehle and Baca (1979). A process related to consistency checking is evaluating the reliability of the data. A hypothesis generated using part of the data for retreival cues may be inconsistent with the remainder of the data for two reasons: The hypothesis may be inappropriate with regards to all of the data or part of the data may be unreliable. Logically, a hypothesis that is inconsistent with an unreliable datum might be worthy of further consideration. This subject specifically recognized that the data might be unreliable. In the following excerpts, Subject D considered the possibility that the battery was the cause of the car not starting, although the car had a recent tune up (Problem 5):

D370: Well, if the battery's dead,

D371: It's an inefficient type guy

D372: Who does it at the station.

D385: This is of course assuming

D386: The mechanic did a fairly decent job.

Ultimately, the subject rejected the battery hypothesis, but reasoned that the generator might be defective.

Subjects' probabilistic responses will be discussed in a following section. However, Subject D's responses revealed that there was some acknowledgement of lack of expertise:

D482: I'm not a mechanical wizard.

D484: Do I look like the Shell Answer Man?

Before making a probability estimate, this subject generally ran over the list of hypotheses and considered their plausibilities one-by-one. Subject D's statements indicated that the hypothesis sets generated were regarded as fairly complete:

D341: I think that is a pretty good possibility.

D342: Those are about them,

D343: I'd say

D344: A pretty high probability.

An apparent pattern in Subject D's protocols was a cycling between reiteration of the data and generation of hypotheses. The hypothesis-generation segments sometimes included a consideration of scenarios and justification of generated hypotheses. The data refreshment phases seemed to serve as intermezzi between bursts of hypothesis-generation activities. The process of considering a scenario, generating a hypothesis and justifying the hypothesis is illustrated in the following excerpts (Problem 4 -- the car stalls at every stop sign):

D1: It could be that the dumb wife does not know how to work the clutch.

D2: So, I think the clutch is a problem.

D42: I feel fairly confident about the clutch.

D43: I've destroyed it myself several times.

The expert subject also appeared to cycle between data rehearsals and generation bursts which were sometimes accompanied by brief scenarios. For example, on Problem 3 where the complaint was that the car was hard to start (see Table 1):

F48: Let me see.

F49: Flooded,

F50: All the time;

F51: Like most of the girls do.

Also, for this subject, deciding how many hypotheses should be generated posed a real problem:

F87: Gonna fill this thing up

F88: With reasons.

F142: Wonder if that's enough.

F143: I don't want him to get upset.

F144: That ought to be enough.

Although subject F could have been estimating probabilities by attending to the substance of the hypotheses generated, the following protocols suggest that a "counting heuristic" may have been employed instead. That is, "a lot" appeared to be functionally related to "very probable":

F100: That'd have to be

F101: At least fifty percent,

F102: If anything.

F103: That's a lotta stuff.

F240: That's a lot of stuff.

F242: I'd say that had to be

F243: At least seventy-five percent

F244: With all that stuff there.

Generated Hypotheses

A speculation having some intuitive appeal is that experts should generate more hypotheses than novices. However, an examination of the frequency with which hypotheses were generated revealed that there is little to distinguish the novice from the expert group on this dependent measure, as illustrated in Table 3.

Insert Table 3 about here

It should be noted that the "frequency" dependent variable is a measure of quantity rather than quality of individual hypotheses. The quality of individual hypotheses can not be assessed in this paradigm. However, the quantity of hypotheses generated can be viewed as a measure of the quality of the collections of hypotheses generated by subjects.

The mean number of hypotheses generated per problem by novice subjects was 3.43 and by expert subjects, 3.36, suggesting that, in lieu of the explicit criteria provided by the experimenter (which was to generate all plausible hypotheses which could be recalled), subjects appeared to adopt the strategy of generating enough hypotheses to fill a "memory span". Although memory span limitations should not have been a factor in the experimental setting, perhaps generating a memory span of hypotheses is the customary strategy of subjects, due to a lifetime of practice.

Deleted from these analyses were responses thought to be inappropriate. For example, one subject suggested that a reason the car refused to start was that it was out of transmission fluid. A hypothesis was judged unacceptable if, in the experimenter's view, it could not have been the proximal cause of the malfunction. By this criterion, seven hypotheses were judged to be unacceptable, which is an average of .1 unacceptable hypotheses per subject per problem.

Table 3 also contains the results of analyzing hypotheses by pooling responses for each problem, accomplished by examining the union of individual subjects' hypothesis sets; that is, examining the set of distinct hypothesis generated by subjects within a group. The mean number of hypotheses in the pooled set, per problem, was 12.6 for novices, 11.2 for experts and 17.8 combined. Thus, on the average, a hypothesis set for one subject on a problem contained 19.2 percent of the distinct hypotheses generated by all subjects on that problem. That is, if the pooled sets of hypotheses for all subjects represent all possible hypotheses, then a typical subject generated less than one-fifth of the possible hypotheses.

An important consideration in comparing the average individual to the pooled group average to obtain the 19.2 percent statistic is the exhaustiveness of the pooled group hypothesis sets. If the pooled hypothesis sets can be shown to be impoverished, then the 19.2 percent statistic would be an underestimate of the proportion of all acceptable hypotheses generated by the average subject.

In the absence of an omniscient automobile mechanic consultant, a permutation analysis was conducted to evaluate the exhaustiveness of the pooled sets. Pooled hypothesis sets were examined for every possible group composed of two subjects to calculate the mean (expected) number of distinct hypotheses in the pooled set. Similarly, all possible pooled sets were examined for groups of each possible size, up to the limit of the total number of subjects. Separate analyses were conducted for the novice group, the expert group and for all subjects combined. Results for the novice and expert groups are listed in Table 4. Figure 1 is a plot of summary results, averaged across problems.

Insert Table 4 and Figure 1 about here

The plots of Figure 1 suggest that adding one more hypothesis generator to a group produces roughly the same enrichment of the pooled set of hypotheses whether the additional subject is an expert or a novice. Also plotted in Figure 1 is a curve representing the slope of the "combined" permutation curve. The slope was calculated for a group of size $\underline{i}$ as the number of hypotheses in the mean pooled set of the group of size $\underline{i}$ minus the number of hypotheses in the mean pooled set of the group of size $\underline{i} - 1$. (The number of hypotheses generated by zero individuals was set at zero.) The point of interest of the "slope" curve is the functional value at the abscissa value of 12, the total number of subjects in the study. This value is related to the exhaustiveness of the pooled set of hypotheses. A slope approaching zero at 12 would indicate that incorporating a 13th subject would not enrich the pooled set. However, as the number of subjects approaches 12, the slope appears to level off at about 1, indicating that an additional subject would be expected to enrich the pool by one hypothesis that was not generated by any of the other subjects.

The number of hypotheses data was also analyzed by employing a simple model of hypothesis generation. To simulate the data plotted in Figure 1, it was supposed that there is a fixed number of hypotheses, $\underline{N}$, available to a group of subjects. Each subject generates a fixed proportion, $\underline{S}$, of those not generated previously (sampling without replacement). Thus, the average subject working individually would generate SN hypotheses. A typical group of two subjects would generate $SN + S\,(N - SN)$ hypotheses, and so on. In other words, in order for a second subject to generate a hypothesis not generated by the first subject, the second subject would need to draw on the pool of hypotheses from which those generated by the first subject had been deleted. The size of this pool for the second subject would be $N - SN$. This recursive description of the model can be represented as a linear differential equation. Letting X symbolize the number of distinct hypotheses generated by a group, Y can be defined as the first derivative of the function relating number of subjects in a group to the corresponding X value. Specifically, Y can be defined for a group of size i as the X value at i minus the X value at $i - 1$. Now, S can be expressed as a function of X, Y and N:

$$S = \frac{Y}{N - (X-Y)} \qquad (1)$$

A couple of elementary algebraic operations are needed to transform Eq. 1 into the following equation:

$$Y = -\left(\frac{S}{1-S}\right)X + \left(\frac{S}{1-S}\right)N \qquad (2)$$

In terms of the parameters of the standard regression equation $Y = mX + b$, the parameters of the model are:

$$S = \frac{m}{m-1} \qquad (3)$$

$$N = -\frac{b}{m} \qquad (4)$$

This model was fitted to the mean data (averaged across problems and across subjects) for the expert, novice and combined groups. Results are given in Table 5.

Insert Table 5 about here

In Table 5, the $\underline{N}$ parameters for the three groups are an estimate of the size of the pool of hypotheses available to the group. This pool is hypothetical -- the actual number of hypotheses generated by the groups were always less than $\underline{N}$. In agreement with intuition, the number of hypotheses in this hypothetical pool, $\underline{N}$, grows with the size of the group.

Also listed in Table 5 are the correlations among the actual values of $\underline{X}$ and the values predicted by applying the definitional recursive representation of the model. That is, for the first $\underline{X}$ value for the novice group, 3.43, the predicted value would be $\underline{S} \times \underline{N} = (.179)(18.1) = 3.24$. Apart from the rather large magnitudes of the correlations, the significant entry in Table 5 is the $\underline{N}$ parameter for the combined group, 21.5. By the yardstick of this model, the combined group of 12 subjects failed to generate (21.5 - 17.8) = 3.7 hypotheses per problem, on the average.

Another indication of the exhaustiveness of the hypothesis sets can be obtained by direct examination of the hypotheses themselves. Table 6 contains all hypotheses generated by subjects for Problem 5.

Insert Table 6 about here

The hypotheses are listed in three categories: those generated by at least one of the novice subjects but by none of the experts, those generated by at least one of the experts but by none of the novices and those generated by at least one novice and one expert subject.

It is difficult but not impossible to generate additional hypotheses for any of the problems. For example, in Problem 5, no subject suggested that the problem could be due to a defective or overlooked after-market anti-theft device in the vehicle. Another possibility is that the starter relay could be defective or the wiring could have been tampered with by a thief in a futile attempt to "hot wire" the car. This consideration and the permutation analysis provide converging evidence in support of the conjecture that pooled hypothesis sets across all 12 subjects are not exhaustive and thus the average subject generated significantly less than one-fifth of all possible hypotheses.

An examination of Table 6 reveals another aspect of the generated hypotheses that was apparent in all problems: the hypotheses generated by the expert subjects seemed to be much more specific than those generated by novices. For example, two experts generated the hypothesis of a defective neutral safety switch, which is highly specific. This hypothesis was not generated by any of the novices. Hypotheses representative of those generated by novices but not by experts include "alternator broken" and "voltage regulator" (defective). Both of these hypotheses are non-specific; a car will start readily with either a broken alternator or a defective voltage regulator; either difficulty would lead to a starting problem only indirectly.

One possible explanation for this pattern is that the experts were able to recall a greater number of possibilities and applied an "it must be specific" criterion to reduce the number of hypotheses to a reasonable number, such as a memory span. Conversely, novice subjects, having less knowledge in their semantic long term stores, would sometimes consider hypotheses only indirectly related to the data in order to generate a comparable number of hypotheses.

Another avenue to account for this pattern of results is to consider each group in terms of the two strategies identified by Hart (1975). Hart termed the strategy of tracing cause and effect patterns in detail to generate hypotheses the "engineering approach". In contrast, the technician relies on experience to suggest likely hypotheses, which are then directly analyzed. Hart commented that when all else fails, the technician is likely to also employ the engineer approach, but only as a last resort. Logical considerations suggest that expert subjects would be inclined to employ a technician approach, while novices would be more likely to employ the engineer approach. An examination of the hypotheses generated by subjects suggested that this was the case; hypotheses generated by experts seemed to be directly linked to the described malfunctions, while novices' responses often could be linked only indirectly to the data, via a causal chain. Presumably, the reason that novices would be more inclined to adopt the "engineer" strategy, tracing out causal links during hypothesis generation, was that their semantic store was not as rich as the typical expert's store. These differing strategies may also help account for the paucity of verbalizations by expert subjects.

Probability Estimates

The mean probability estimate was 69.2 percent for novices (range: 17 to 98) and 67.5 for experts (range: 27 to 100). A significant problem in evaluating subjects' probabilistic estimates is the unavailability of veridical values, which have proven useful in demonstrating that subjective estimates were typically excessive in similar contexts (e.g. see Mehle, Gettys, Manning, Baca and Fisher, 1979). In an attempt to establish that estimates obtained in this study were excessive, an analysis technique dubbed the "they can't all be right" procedure was devised. This procedure consists of examining the hypotheses generated by subjects and performing permissible operations under the (temporary) assumption that subjects estimates are consistent with the axioms of probability theory.

To illustrate, suppose that one subject generates only two hypotheses (battery and regulator) and estimates the probability of this set as 80 percent. Suppose a second subject generates only one hypothesis (battery) and estimates its probability as 50 percent. Assuming that the hypotheses are mutually exclusive, (a reasonable assumption in this context), a permissible inference is that the probability of the hypothesis "regulator" is 80 - 50, or 30 percent. Working in this manner, it is possible to obtain a collective estimate for the probability of the pooled (over all 12 subjects) set of hypotheses for a problem. These estimates are listed in Table 7 as the "unadjusted estimates".

Insert Table 7 about here

If this collective estimate is in excess of 100 percent, the conclusion would be that the collection of subjects' estimates are not in agreement with the probability theory axioms. In particular, collective estimates well in excess of 100 percent suggest strongly that a typical subject would be excessive. Such a result would not permit an identification of any particular subject as extreme; rather, it would lead to a characterization of the typical subject as extreme.

One problem with this approach is that, due to the pattern of subjects' responses, the collective estimates are for proper subsets of the pooled sets, that is, no estimates can be made for some elements of the pooled sets. It seems reasonable to assume that there are no intrinsic differences between hypotheses included in the collective estimate and those excluded. (This assumption may be suspect, but it is not really crucial to the conclusion). To compensate for this difficulty, estimates were adjusted by simply multiplying by the number of hypotheses in the pooled set and dividing by the number of hypotheses used to compute the unadjusted estimate. This adjustment is equivalent to estimating the probability of hypotheses excluded from the collective estimate as the mean of those included in the collective estimate.

These "adjusted" estimates are also listed in Table 7. Both the adjusted and the unadjusted estimates support the conclusion that subjects "could not all have been right". The typical subject was excessive in assessing the probability of generated hypotheses. For example, the mean adjusted estimate of the collective set is 504 percent, which is clearly in excess of 100 percent. It should be noted that this adjusted estimate is also somewhat of an underestimate. The pooled set of all hypotheses generated by subjects is a proper subset of the set of all acceptable hypotheses. The pooled set is incomplete for reasons discussed in the previous "Generated Hypotheses" section.

Summary

The main results of the protocol analyses included the findings that subjects explictly considered hypothesis consistency and data reliability during hypothesis generation. While occasionally acknowledging lack of expertise, subjects generally regarded their hypothesis sets as fairly exhaustive. The patterns of the protocols suggested that hypotheses were generated in bursts, sometimes accompanied by the construction of plausible scenarios.

An analysis of the generated hypotheses demonstrated that subjects generated about 3.4 hypotheses per problem, regardless of whether they were experts or novices. A permutation analysis and content considerations led to the conclusion that hypothesis sets obtained by pooling the responses of all subjects were incomplete. Typical subjects generated less that one-fifth of the acceptable hypotheses for a problem, while regarding their generated sets as fairly probable. Analyses of the probabilistic responses yielded a conclusion that subjects were typically quite excessive in their estimates.

Taken together, the results of these analyses lead to a rather discouraging characterization of the typical hypothesis generator in this study. The typical subject generated quite impoverished sets of hypotheses, yet were excessive in estimating the exhaustiveness of their hypothesis sets. If low perceived plausibility of the hypothesis set does serve as the stimulus for resumption of hypothesis generation activities (Gettys and Fisher, 1979), then subjects do not generate hypotheses when they should in real-world problem-solving situations. It is clearly not optimal, working with a limited information-processing system, for subjects to always carry an exhaustive set of hypotheses through the decision-making process, particularly when the number of hypotheses in an exhaustive set is extremely large. However, in applied settings, there exists a large class of decision problems which require the decision maker to generate exhaustive, or nearly exhaustive, hypothesis sets -- for example, in nuclear power and medical decision making. In such problems, generating less than one fifth of the possible hypotheses may be very costly. Encouraging decision makers to, for example, make use of an artificial memory aid to enrich the set of hypotheses considered, holds promise for significantly improving the entire decision process.

References

Bhaskar, R. and Simon, H. A. Problem solving in semantically rich domains: an example from engineering thermodynamics. _Cognitive Science_, 1977, _1_, 193-215.

Brooks, R. Towards a theory of the cognitive processes in computer programming. _International Journal of Man-Machine Studies_, 1977, _9_, 737-751.

Churchman, C. W. and Buchanan, B. G. On the design of inductive systems: Some philosphical problems. _British Journal of the Philosophy of Science_, 1969, _20_, 311-323.

Ericsson, K. A. and Simon, H. A. Retrospective verbal reports as data. Carnegie-Mellon University C.I.P. Working Paper No. 388, 1978.

Fischhoff, B., Slovic, P. & Lichtenstein, S. Knowing with certainty: The appropriateness of extreme confidence. _Journal of Experimental Psychology: Human Perception and Performance_, 1977, _3_, 552-564.

Fischhoff, B., Slovic, P. & Lichtenstein, S. Fault trees: Sensitivity of estimated failure probabilities to problem representation. _Journal of Experimental Psychology: Human Perception and Performance_, 1978, _4_, 330-344.

Fisher, S. D., Gett, J. C. F., Manning, C., Mehle, T. & Baca, S. Consistency checking in hypothesis generation. University of Oklahoma Decision Processes Laboratory Technical Report TR 29-7-79, 1979.

Gerwin, D. Information processing, data inferences and scientific generalization. _Behavioral Science_, 1974, _19_, 314-325.

Gerwin, D. and Newsted, P. A comparison of some inductive inference models. _Behavioral Science_, 1977, _22_, 1-11.

Gettys, C. F. and Fisher, S. D. Hypothesis plausibility and hypothesis generation. Organizational Behavior and Human Performance, 1979, 24, 93-110.

Gettys, C. F., Mehle, T., Baca, S., Fisher, S. & Manning, C. A memory retrieval aid for hypothesis generation. University of Oklahoma Decision Processes Laboratory Technical Report TR 27-7-79, 1979.

Hanson, N. R. Is there a logic of scientific discovery? In H. Feigl and G. Maxwell (Eds.), Current Issues in the Philosophy of Science. New York: Holt, 1961, 20-35.

Hart, P. E. Progress on a computer based consultant. Stanford Research Institute Artificial Intelligence Group Technical Note 99, 1975.

Lichtenstein, S., Fischhoff, B. and Phillips, L. D. Calibration of probabilities: The state of the art. In H. Jungermann and G. deZeeuw (Eds.), Decision Making and Change in Human Affairs. Dordrecht, Holland: D. Reidel, 275-324.

Mehle, T., Gettys, C. F., Manning, C., Baca, S. and Fisher, S. The availability explanation of excessive plausibility assessments. University of Oklahoma Decision Processes Laboratory Technical Report TR 30-7-79, 1979.

Milton, J. (Ed.) Chilton's Foreign Car Repair Manual, Volume I. Radnor, Penn.: Chilton Book Company, 1971.

Nisbett, R. E. and Wilson, T. D. Telling more than we can know: verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Payne, J. W. Task complexity and contingent processing in decision making: An information search and protocal analysis. Organizational Behavior and Human Performance, 1976, 16, 366-387.

Payne, J. W., Braunstein, M. L. & Carrol, J. S. Exploring predecisional behavior: An alternative approach to decision research. Organizational Behavior and Human Performance, 1978, 22, 17-44.

Rouse, W. B. Human problem solving performance in a fault diagnosis task. IEEE Transactions on Systems, Man and Cybernetics, 1978, 8, 258-271. (a)

Rouse, W. B. A model of human decisionmaking in a fault diagnosis task. IEEE Transactions on Systems, Man and Cybernetics, 1978, 8, 357-361. (b)

Sacerdotti, E. D. The nonlinear nature of plans. Stanford Research Institute Artificial Intelligence Group Technical Note 101, 1975.

Shanteau, J. and McClelland, G. H. Mental search processes in problem solving. Memory and Cognition, 1975, 3, 627-634.

Slovic, P. and Fischhoff, B. On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 544-551.

Simon, H. A. The functional equivalence of problem solving skills. Cognitive Psychology, 1975, 7, 268-288.

Simon, H. A. Information-processing theory of human problem solving. In W. K. Estes (Ed.) Handbook of Learning and Cognitive Processes, Volume 5: Human Information Processing. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Simon, H. A. Information processing models of cognition. Annual Review of Psychology, 1979, 30, 363-396.

Tversky, A. and Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 4, 207-232.

Wortman, P. M. Representation and strategy in diagnostic problem solving. Human Factors, 1966, 8, 48-53.

Wortman, P. M. Cognitive utilization of probabilistic cues. Behavior Science, 1970, 15, 329-336.

Wortman, P. M. Medical diagnosis: An information-processing approach. Computers and Biomedical Research, 1972, 5, 315-328.

Wortman, P. M. and Kleinmuntz, B. The role of memory in information-processing models of problem solving. Unpublished manuscript, 1973.

Zadeh, I. A. Fuzzy sets. Information and Control, 1965, 8, 338-353.

Footnotes

This research was based on a Ph.D. dissertation undertaken at the University of Oklahoma, Norman, supervised by Charles F. Gettys. The author is grateful for the support and many constructive comments and suggestions of Professor Gettys. Also, the article has benefited from the comments of the other dissertation committee members, Raymond Dacey, W. Alan Nicewander, Steven M. Smith and Larry Toothaker. The assistance and helpful comments of Suzanne Baca, Carol Manning and Stan Fisher were also much appreciated. The research was supported in part by a contract from the Office of Naval Research.

Requests for reprints should be sent to Tom Mehle, Department of Psychology, 209 Burnett Hall, University of Nebraska - Lincoln, Lincoln NE 68588.

[1]The inspiration for this analysis was a suggestion by Charles Gettys.

Table 1

Problem Stimuli

Number	Problem Stimuli
1	THE CAR IS AMERICAN WITH AN EIGHT CYLINDER ENGINE AND AN AUTOMATIC TRANSMISSION; IT IS TWO YEARS OLD AND IS DUE FOR A TUNE UP. THE PROBLEM IS THAT THE CAR REFUSES TO START. THE ENGINE TURNS OVER AND THERE IS A GAS SMELL
2	THE CAR HAS A MANUAL TRANSMISSION AND A SIX CYLINDER ENGINE. IT IS AN IMPORTED MODEL AND IS LESS THAN A YEAR OLD; IT HAS HAD A RECENT TUNE UP. YOUR SPOUSE COMPLAINED THAT ALTHOUGH THE CAR STARTS FINE, IT IS MAKING STRANGE NOISES. ALSO, BOTH THE 'OIL' AND THE 'HOT' WARNING LIGHTS CAME ON WHILE DRIVING BACK FROM A SHOPPING TRIP.
3	THE CAR IS AMERICAN WITH A FOUR CYLINDER ENGINE AND AN AUTOMATIC TRANSMISSION. THE CAR IS FIVE YEARS OLD AND IS IN NEED OF A TUNE UP. THE CAR TROUBLE MENTIONED BY YOUR SPOUSE WAS THAT THE CAR IS HARD TO START AND THE 'HOT' WARNING LIGHT COMES ON WHEN THE CAR IS DRIVEN FOR ANY LENGTH OF TIME.
4	THE CAR IS A FOREIGN FOUR-CYLINDER MODEL WITH A MANUAL TRANSMISSION. IT HAS HAD A TUNE-UP RECENTLY AND IS LESS THAN TWO YEARS OLD. THE PROBLEM WITH THIS CAR IS THAT THE ENGINE HAS A TENDENCY TO DIE AT EVERY STOP SIGN AND STOP LIGHT. THE CAR STARTS FINE AND NO WARNING LIGHTS ARE COMING ON.
5	YOUR CAR IS SEVEN YEARS OLD AND IS AN AMERICAN SIX-CYLINDER MODEL. IT HAS AN AUTOMATIC TRANSMISSION AND HAS HAD A TUNE-UP RECENTLY. YOUR SPOUSE COMPLAINED THAT THE CAR WOULD NOT START -- IT WAS TOTALLY DEAD. THERE WAS NOT EVEN A CLICK WHEN THE KEY WAS TURNED.

Table 2

Protocol Frequencies

	Novices		Experts	
	Subject Letter	Total Number of Protocols	Subject Letter	Total Number of Protocols
	A	74	G	244
	B	70	H	36
	C	87	I	68
	D	539	J	48
	E	177	K	60
	F	55	L	13
Mean per Problem		33.4		15.6
Mean per Subject		167		78.2
Median per Subject		80.5		54

Table 3

Hypothesis Frequencies

Problem Number	Novice Subjects: Mean Number of Hypotheses per Subject	Novice Subjects: Number in Pooled Set	Expert Subjects: Mean Number of Hypotheses per Subject	Expert Subjects: Number in Pooled Set
1	3.17	14	3.17	10
2	2.83	11	2.83	8
3	4.67	19	4.00	17
4	3.17	9	2.67	10
5	3.33	10	4.17	11
Mean	3.43	12.6	3.36	11.2

Novice and Expert Subjects (Pooled)

Problem Number	Number of Unacceptable Hypotheses	Mean Number of Hypotheses per Subject	Number in Pooled Set
1	2	3.17	18
2	1	2.83	14
3	1	4.33	28
4	2	2.92	15
5	1	3.75	14
Mean	1.4	3.40	17.8

Table 4

Permutation Analysis

Mean Number of Hypotheses in Pooled Groups

Problem Number	Number in Pooled Group: 1	2	3	4	5	6
Novices						
1	3.2	6.0	8.5	10.7	12.5	14.0
2	2.8	5.1	6.9	8.5	9.8	11.0
3	4.7	8.0	11.1	13.9	16.5	19.0
4	3.2	5.0	6.3	7.3	8.2	9.0
5	3.3	5.5	7.1	8.2	9.2	10.0
Experts						
1	3.2	5.4	7.0	8.2	9.2	10.0
2	2.8	4.6	5.8	6.6	7.3	8.0
3	4.0	7.2	10.1	12.9	15.5	18.0
4	2.7	4.3	5.8	7.3	8.7	10.0
5	4.2	6.1	7.6	8.9	10.0	11.0

Table 5

Model Fitting Results

Group	Sampling Rate (S)	Number of Hypotheses (N)	Correlation of X to X
Novice	.179	18.1	.999
Expert	.196	15.5	.998
Combined	.134	21.5	.998

Table 6

Hypotheses Generated on Problem 5

Hypotheses Generated By At Least One:		
Novice but no Experts	Expert but no Novices	Novice and One Expert
Alternator broken	Neutral safety switch	Battery cables broken
Mechanical breakage	Ignition switch	Battery terminals
Voltage regulator	Stolen Motor	Starter
	Not in 'P' or 'N'	Ignition
		Slipping belt
		Solenoid
		Battery

Table 7

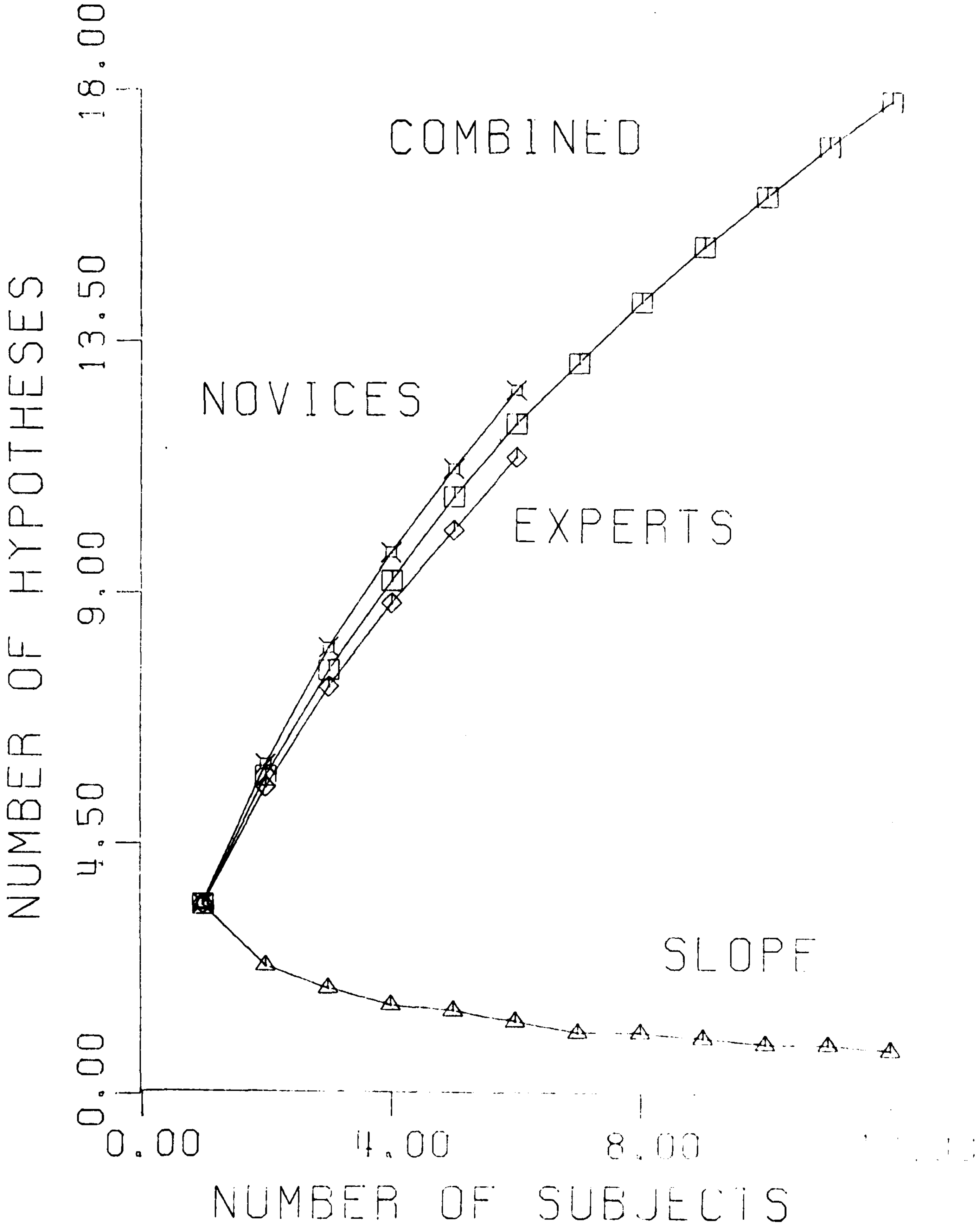
Collective Estimates of Hypothesis

Set Probabilities (Percent)

	Problem Number					
Type of Estimate	1	2	3	4	5	Mean
Unadjusted Estimate	301	250	269	263	199	256
Adjusted Estimate	542	389	628	564	398	504

Figure Caption

Plotted are the results of a permutation analysis, averaged across five problems. The "combined" curve was obtained by pooling all 12 subjects in the study. The "slope" curve is the rate of change of the "combined" curve. If the slope is not effectively zero at 12 subjects, then the pooled set of hypotheses over 12 subjects could be regarded as incomplete.



COMBINED
NOVICES
EXPERTS
SLOPF
NUMBER OF HYPOTHESES
18.00
13.50
9.00
4.50
0.00
0.00
4.00
8.00
NUMBER OF SUBJECTS

OFFICE OF NAVAL RESEARCH, CODE 455
TECHNICAL REPORTS DISTRIBUTION LIST

CDR Paul R. Chatelier
Office of the Deputy Under Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217 (5 cys)

Director
Analysis and Support Division
Code 230
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director Statistics and Probability
Program
Code 436
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director
Information Systems Program
Code 437
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Commanding Officer
ONR Branch Office
ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, Massachusetts 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
536 South Clark Street
Chicago, Illinois 60605

Commanding Officer
ONR Branch Office
ATTN: Mr. R. Lawson
1030 East Green Street
Pasadena, California 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, California 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Dr. Robert G. Smith
Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, Florida 32813

ONR, Code 455, Technical Reports Distribution List

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-00T
Orlando, Florida 32813

Scientific Advisor to DCNO (MPT)
OP 01T (Dr. Marshall)
Washington, D.C. 20370

CDR Thomas Berghage
Naval Health Research Center
San Diego, California 92152

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, Connecticut 06340

Navy Personnel Research and Development Center
Manned Systems Design, Code 311
San Diego, California 92152

Navy Personnel Research and Development Center
Code 305
San Diego, California 92152

Navy Personnel Research and Development Center
Management Support Department
Code 210
San Diego, California 92151

CDR P. M. Curran
Code 604
Human Factors Engineering Division
Naval Air Development Center
Warminster, Pennsylvania 18974

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, Maryland 21402

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, California 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, California 93940

Mr. Warren Lewis
Human Engineering Branch
Code 8231
Naval Ocean Systems Center
San Diego, California 92152

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 08D22
Washington, D.C. 20360

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D.C. 20362

Naval Sea Systems Command
Personnel & Training Analyses Office
NAVSEA 074C1
Washington, D.C. 20362

LCDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, California 93940

Mr. Merlin Malehorn
Office of the Chief of Naval Operations (OP 102)
Washington, D.C. 20350

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, D.C. 20310

ONR, Code 455, Technical Reports Distribution List

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Dr. Edgar M. Johnson
Organizations and Systems Research
Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

U.S. Army Aeromedical Research Lab
ATTN: CPT Gerald P. Krueger
Ft. Rucker, Alabama 36362

ARI Field Unit-USAREUR
ATTN: Library
C/O ODCSPER
HQ USAREUR & 7th Army
APO New York 09403

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, Ohio 45433

Air University Library
Maxwell Air Force Base, Alabama 36112

Dr. Gordon Eckstrand
AFHRL/ASM
Wright-Patterson AFB, Ohio 45433

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
Establishment
Teddington, Middlesex TW11 OLN
ENGLAND

Director, Human Factors Wing
Defense & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, Virginia 22314 (12 cys)

Dr. Craig Fields
Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, Virginia 22209

Dr. Judith Daly
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, Virginia 22209

ONR, Code 455, Technical Reports Distribution List

Dr. Stanley Deutsch
Office of Life Sciences
National Aeronautics and Space
  Administration
600 Independence Avenue
Washington, D.C. 20546

Professor Douglas E. Hunter
Defense Intelligence School
Washington, D.C. 20374

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, California 93017

Dr. Gary McClelland
Institute of Behavioral Sciences
University of Colorado
Boulder, Colorado 80309

Human Resources Research Office
300 N. Washington Street
Alexandria, Virginia 22314

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Melo Park, California 94025

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

Professor Judea Pearl
Engineering Systems Department
University of California-Los Angeles
405 Hilgard Avenue
Los Angeles, California 90024

Professor Howard Raiffa
Graduate School of Business
  Administration
Harvard University
Soldiers Field Road
Boston, Massachusetts 02163

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, Pennsylvania 19087

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, Oregon 97401

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, California 94305

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, California 91364

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, NW
Washington, D.C. 20036

Dr. Ward Edwards
Director, Social Science Research
  Institute
Unversity of Southern California
Los Angeles, California 90007

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Room 201
Boulder, Colorado 80309

Dr. William Howell
Department of Psychology
Rice University
Houston, Texas 77001

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, N.W.
Washington, D.C. 20036 (3 cys)

Dr. Clinton Kelly
Decisions and Designs, Inc.
8400 Westpark Drive, Suite 600
P.O. Box 907
McLean, Virginia 22101

Mr. Richard J. Heuer, Jr.
27585 Via Sereno
Carmel, California 93923

ONR, Code 455, Technical Reports Distribution List

Mr. Tim Gilbert
The MITRE Corporation
1820 Dolly Madison Blvd.
McLean, Virginia 22102